CUE-RESPONSE SEPARATION AND ELEMENT PROXIMITY IN THE FEATURE DISCRIMINATION PARADIGM

An abstract of a Thesis by
Thomas B. Umphress
December 1975
Drake University
Advisor: William D. Klipec

The problem. To investigate the function of the proximity of elements in a feature discrimination paradigm on the feature-positive effect, remote responding, and stimulus identification.

Procedure. Sixty preschoolers between the ages of three and five years were trained to discriminate between two simultaneously presented displays containing either four common features, or three common and one distinctive feature. One-third of the subjects were reinforced for touching any feature on the display with the distinctive feature (FP-1 group), one-third for touching a common feature on the distinctive feature display (FP-2 group), and one-third for touching any feature on the common feature display (FN group). Half of each group of subjects were trained with the features compacted in the centers of the displays and half with the features distributed in the outer corners of the displays.

Findings. A clear feature-positive effect occurred only when the elements were distributed. Compacting the elements resulted in fewer errors but did not eliminate a significant difference between the FP-2 and FN group. Generalization tests given during extinction indicated that subjects responded on the basis of features, pattern, or both independent of the condition to which they were assigned.

Conclusions. Responses to features were influenced by reinforcement probability but responses did not have to be directed to a feature for it to affect responding. The hypothesis that cue-response separation contributes to feature-negative difficulty was not supported.

Recommendations. Future research in this area should attempt to clarify the parameters of the effect, remote responding strategies, and factors that affect responding.

CUE-RESPONSE SEPARATION AND ELEMENT PROXIMITY IN THE FEATURE DISCRIMINATION PARADIGM

A Thesis

Presented to

The School of Graduate Studies

Drake University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

---

by

Thomas B. Umphress

December 1975

1975
Umph

CUE-RESPONSE SEPARATION AND ELEMENT PROXIMITY IN THE FEATURE DISCRIMINATION PARADIGM

by

Thomas B. Umphress

Approved by Committee:

William D. Klipec
Chairman

W. Scott Wood

Joseph A. Fisher

Earle L. Canfield
Dean of the School of Graduate Studies

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

"Discrimination learning is the name we give to the process by which stimuli come to acquire selective control over behavior (Hilgard & Bower, 1966, p. 521)." Discrimination learning has been an important tool in the investigation of learning phenomena because it seems to be involved in all learning to some degree.

One common paradigm through which organisms come to discriminate among stimuli is differential reinforcement. In a typical differential reinforcement training paradigm, one stimulus is arbitrarily designated as the positive or S+ stimulus and responses in its presence are followed by reinforcement while responses in the presence of another stimulus (designated as the negative or S- stimulus) are not reinforced. Operationally, discrimination training is a combination or conditioning and extinction procedure, where the S+ trials are conditioning trials and the S- trials are extinction trials. Discrimination training increases the probability of a response occuring in the presence of the S+ stimulus and reduces the probability of a response occuring in the presence of the S- stimulus. In a successive discrimination only one stimulus is presented at a time with stimuli separated in a temporal sequence. In a simultaneous discrimination the S+ and S- stimuli are presented together and the organism can respond to only one of the spatially separate displays.

In a feature discrimination paradigm the discrimination is between stimulus displays which are differentiated only by the presence of a single distinctive feature. That is, one display contains only _A_ elements while the other display contains a _B_ element in addition to the _A_ elements. In the feature-positive case, responses to the _AB_ display are reinforced and responses to the _A_ display are not reinforced. In the feature-negative case, the reinforcement contingency is reversed. The _A_ display is the S+, or positive display, and the _AB_ display is the S-, or negative display. Thus, the distinction between the feature-positive and feature-negative conditions is determined by the location of the distinctive feature on the positive or negative display.

In a simple differential conditioning paradigm, if two stimuli differ along one dimension (e.g., vertical vs. horizontal line), the acquisition of the two-choice discrimination is not usually influenced by the assignment of the stimuli to the positive and negative conditions. This is not the case, however, in the feature discrimination paradigm. Previous work has shown that if two stimulus displays were differentiated by a single distinctive feature, locating the distinctive feature on the positive display facilitated learning more than if the distinctive feature were on the negative display. In this case, acquisition of the discrimination is not symmetrical. When the distinctive feature is located on the S+ display, the discrimination is made with fewer errors than if it is located on the negative display (Jenkins & Sainsbury, 1969, 1970; Sainsbury, 1971a, 1971b, 1973). This phenomenon has been termed the feature-positive effect by Jenkins and Sainsbury (1969, 1970) and has been found with all the stimulus arrangements in Figure 1. The effect is most likely to occur when the two stimulus displays differ asymmetrically. Jenkins and Sainsbury (1970) suggest a principle to be used in determining whether a pair of stimulus displays are symmetrical or asymmetrical. They suggest that if all the features common to the two displays are removed, and nothing remains, then the displays are symmetrical: if, however, that operation leaves one display with something remaining, then the two displays are asymmetrical.

The feature-positive effect has been demonstrated with pigeons using a successive discrimination (Jenkins & Sainsbury, 1969, 1970; Sainsbury, 1971a). In this go/no-go type discrimination, the pigeons learned to respond differentially to the two displays in the feature-positive case but not in the feature-negative case. In the feature-positive case, pigeons eventually directed their responses on the S+ display to the *B* feature before ceasing to

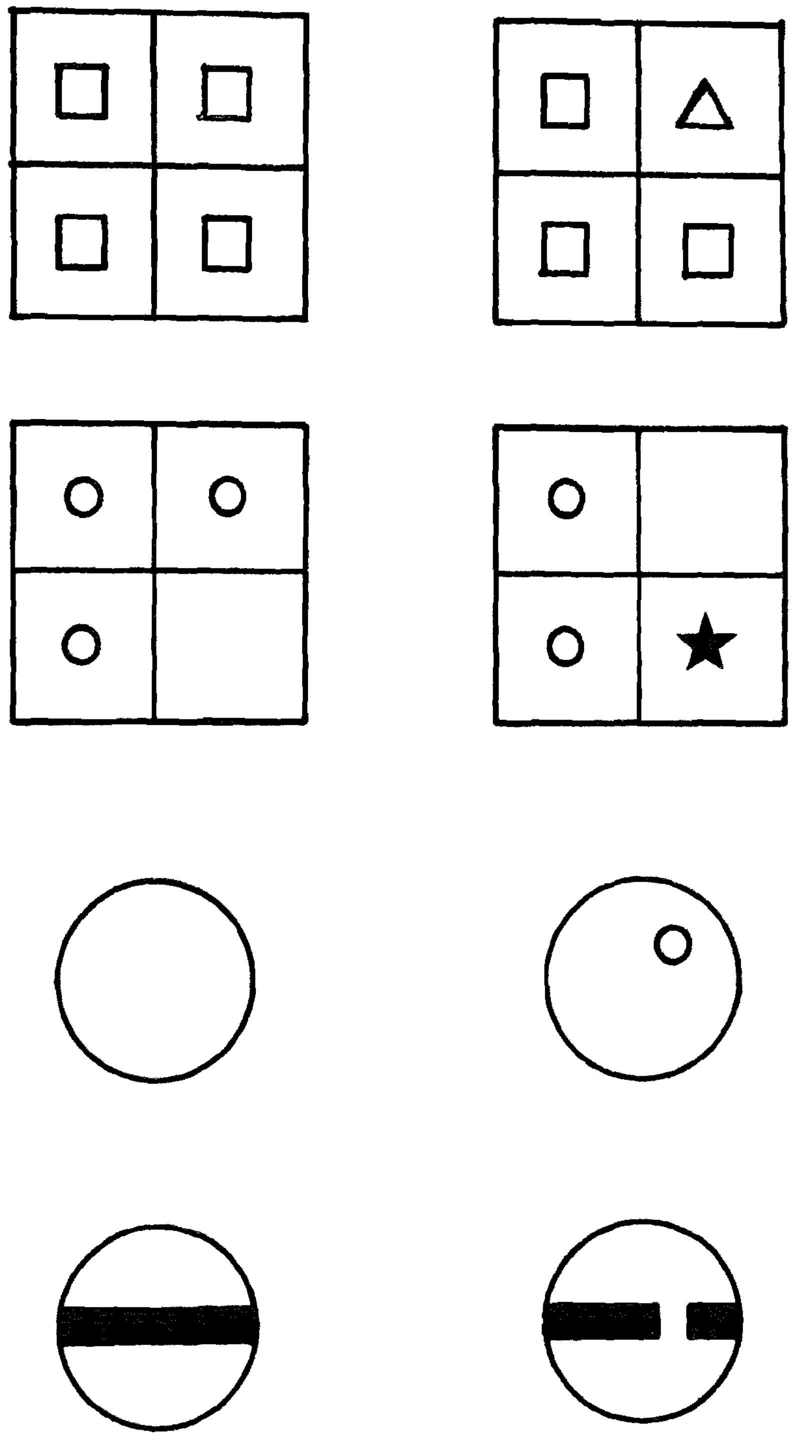

Fig. 1. Typical feature discrimination displays. The common feature display is on the left and the distinctive feature display is on the right.

respond to the S- display. In the feature-negative case, the pigeons failed to withhold responding to the negative display AB even though they did not peck B.

Sainsbury (1971a) used a stimulus complex that was divided into four quadrants. When only one element occupied each quadrant, pigeons failed to learn the discrimination in the feature-negative condition, but when all four elements were compacted into one quadrant the feature-negative discrimination was facilitated. One explanation of Sainsbury's results is based upon a Gestalt interpretation that states that bringing the elements closer together created patterns to which the subject could respond. Thus, the discrimination became a simple choice discrimination between two patterns. Sainsbury suggests that for learning to take place in the feature-negative condition, common and distinctive elements within the AB display must be discriminated from each other and that compacting the display may provide the necessary interaction by allowing B to suppress AB responses. The results of responses to test displays during extinction (Sainsbury, 1971) suggest that in the feature-negative case, the pattern created by the compacted display exerted a great deal of control over responding. However, one test display could show no evidence of responses to common features being inhibited by the presence of a distinctive feature.

It should be pointed out that Jenkins and Sainsbury

(1969) use the concept of inhibition strictly in an operational sense when they state that, "We do not know whether the distinctive feature on the negative display exerts inhibitory control. The observation is simply that the locus of responding shifts away from it to common features." Other theorists, notably Hull (1943) and Spense (1952) deal with inhibition as an intervening variable with surplus meaning, thus depicting it as more of a process than as simply the result of the manipulation of empirical variables.

Using Pigeons as subjects, Hearst (1969) reported superior learning in the feature-negative condition. He used two circular displays that were differentiated by the presence of a straight line on one of the displays. Hearst initially proposed that the blank key and the keys with different line tilt values were orthogonal. He suggested that one stimulus value was equidistant from the other and that learning should be the same for either condition. Hearst also noted basic procedural differences between his experiments and the pigeon study of Jenkins and Sainsbury (1969). Hearst used a VI schedule of reinforcement and the operant chamber was blacked out during intertrial intervals. Jenkins and Sainsbury used a limited trial procedure and the chamber was continuously illuminated. Their subjects had to peck four times within seven seconds or wait until the next trial to receive reinforcement. Jenkins and

Sainsbury gave non-differential pretraining while Hearst pretrained his birds to peck the blank key. When Hearst introduced the line in phase I, the blank key should have had an initial advantage because of its history of reinforcement. If after pretraining, responding was under control of the blank key, learning would be faster in the feature-negative condition because the blank key was the positive stimulus. In the feature-positive condition the organism's initial preference to the blank key would have to be extinguished before the line could gain control over responding.

Norton, Muldrew, and Strub (1971) found the feature-positive effect with children using a simultaneous discrimination and with adults using a successive discrimination. With humans this effect is influenced by maturation. Sainsbury (1973) found this effect to gradually disappear within older age groups until each condition was learned equally easily. Sainsbury found the effect for all his 4 year-old group, for half of his 7 year-old group, and for none of his 9 year-old group. The youngest group, unlike pigeons failed to withhold responding to B on the S- display. When errors to criterion was used as the dependent measure, the feature-positive effect was very evident in the 5 and 7 year-old groups. The number of errors to criterion declined with age.

With children, Pilek (1972) found that the feature-positive effect could be eliminated with prior training on easier feature-negative discriminations involving fewer elements. The Pilek (1972) experiment involved two displays which totaled either 3, 5, or 7 common elements. Five year-olds were unable to learn in the feature-negative condition with 7 common elements unless they were first trained to criterion with the 3 and 5 common element displays. Bitgood, Segrave, and Jenkins (Note 1) found that consequating responses to the S- display in the feature-negative condition with mild verbal punishment facilitated learning the discrimination more than a neutral or blank consequence. This effect was attributed to misinterpretation of the lack of explicit consequences for an incorrect response.

Whenever a phenomena such as the feature-positive effect is discovered, existing theories of learning are examined to see if they can account for the results. Not all theories of learning agree on how an organism establishes a discrimination. An examination of some of these theories will show how the feature-positive effect relates to various learning theories.

Theories of discrimination learning can be categorized as continuity theory or non-continuity theory. The continuity theory emphasizes the gradual accumulation of habit strength and the interaction of gradients of excitation and inhibition. According to this type of theory organisms base their discriminations upon absolute stimulus values and they learn equally about all cues impinging upon their receptors. This type of theory has also been referred to as the nonselection, absolute, or nonattention theory. The noncontinuity theory emphasizes active problem solving by an organism which tries out different hypotheses to solve the discrimination. Noncontinuity theory implies that the organism bases his responses on relationships among stimuli. This type of theory has also been referred to as selection, relational or attention theory.

One of the first phenomena to test continuity and noncontinuity theory was transposition. This is the apparent ability of an organism to make a discrimination based upon a relationship rather than an absolute stimulus value. This phenomena is illustrated in experiments by Spence (1937a, 1937b). Spence trained chimpanzees to respond to the larger of two circles. Later, the chimps were presented with the original reinforced circle and one that was larger. The chimps chose the larger circle over the previously reinforced circle. The chimps apparently were responding to the concept of "larger than" instead of responding to a specific size. Spence incorporated the notion of algebraic summation within a gradient of generalization to account for transposition. He hypothesized that response strength was greatest not at the original S+ value but at a larger stimulus value. This explanation would temporarily strengthen one weakness of continuity theory.

The reversing of discriminative cues in the presolution period is one method of investigating continuity and noncontinuity theories. While the subject is still responding at chance level, S+ and S- are reversed. This procedure is known as the reversal shift. If the subject is using a hypothesis during presolution and has not demonstrated learning, reversing the cues should not effect learning because the subject was using the wrong hypothesis. If, however, response strength and response inhibition are being accumulated on S+ and S- trials as Hull (1943) or Spence (1952) suggest, then a reversal shift should hinder learning. The habit strength accumulated to the previous S+ stimulus would first have to be extinguished before inhibition strength could be accumulated. The inhibition accumulated to the previous S- would first have to be overcome before habit strength could start to accumulate. Continuity theory would propose an initial disadvantage to the subject when cues are reversed. Ehrenfreund (1948) and Mackintosh (1965b) have conducted experiments which demonstrate that reversing cues does retard learning. Reid (1953), Pubols (1956), and Sperling (1965a, 1965b) have, on the other hand, found that when cues were reversed after the discrimination was well established made learning the reversal easier. There also have been experiments reported in which these results were not obtained (D'Amato & Schiff, 1965).

Continuity theory states that the organism makes an absolute discrimination and noncontinuity theory states that the organism makes a relational discrimination. If the organism does solve the discrimination based on a relationship, then the opportunity to compare stimuli as in the case of a simultaneous discrimination should be easier than a successive discrimination where no direct comparison of stimuli is available. Because one method of presenting stimuli involves a choice and the other does not, a direct comparison of the methods is difficult. A classic experiment by Lawrence and DeRivera (1954) did investigate this question. Rats were given a discrimination which they could solve on the basis of a relationship between stimuli or on the basis of absolute stimulus values. Transposition tests administered after learning showed that about 80% of the rats were responding according to a relational interpretation and 20% according to an absolute interpretation.

"Attention-like" phenomena have also been used in attempts to resolve the issue of continuity versus noncontinuity. The fact that compound stimuli are not equally conditionable has contributed to the acceptance of the notion that mere continuity is not a sufficient condition for learning. Currently, there is no single theory that can account adequately for all of the phenomena associated with discrimination learning.

Traditional theories of discrimination are of the continuity type. Since discrimination training is a combination of conditioning and extinction procedures, many theories have made use of these operations to explain discrimination learning. Spence (1952) and Hull (1943) represent the traditional conditioning-extinction theorists. The following assumptions are incorporated into their theories: (Kimble, 1961, p. 364)

1. that every reinforcement leads to an increment in the (excitatory) tendency to repeat a response.
2. that every non-reinforcement leads to an increment in the (inhibitory) tendency not to respond.
3. that both of these tendencies generalize to other stimuli.
4. that the magnitude of the inhibitory tendency is less than that of the excitatory tendency.
5. that the excitatory and the inhibitory tendencies interact algebraically.
6. that discriminatory reactions are based on their resolution of the competing tendencies in favor of the reaction to the stimulus which has the stronger tendency conditioned (or generalized) to it.

Spence and Hull proposed that a gradient of excitation associated with reinforcement and a gradient of inhibition associated with extinction interact algebraically to produce varying degrees of response strength. This concept of algebraic summation was adopted to account for response strength. Response strength is represented in the following equation:

$$R_a = R - I$$

R_a represents net response strength

R represents strength of generalization of reinforcement

I represents strength of generalization of extinction

The net strength of a response at any moment is the strength of the generalization of reinforcement minus the strength of the generalization of extinction.

Jenkins and Sainsbury (1969, 1970) developed a simultaneous discrimination theory to interpret data generated from the feature discrimination paradigm. This theory proposes a simultaneous discrimination within the elements of the feature display and between the two displays. The displays are viewed as being composed of two types of elements--common (_A_) and distinctive (_B_). The subject may respond to any element rather than to the display as a whole. In the feature-positive (FP) arrangement, the discrimination is an _AB_+_A_- type (the slanted line indicates a separation of displays and "+" indicates the positive display and "-" the negative display). The feature-negative (FN) arrangement is an _A_+_AB_- discrimination. In the FP case, the probability (P) of a response to _B_ being reinforced is 1.0 while the P(A) = .50. On the _AB_+ display a simultaneous discrimination exists between the _A_ elements and the _B_ element. The element _B_ eventually gains control over the response because of the greater reinforcement probability. Thus, all elements of the display are considered independent elements. In the FP case, an organism needs only to learn to respond to _B_ and not to _A_. No interaction is necessary. As the theory predicts, subjects localize their responses to B prior to their termination of responding to the negative display (Sainsbury, 1973). In the RN case, A responses have the same probability of reinforcement as the FP case, initially, but now P(B) = 0. In the RN case some interaction between elements is necessary if the discrimination is to be learned. A conditional discrimination must develop. The subject must learn to respond to A unless B is present. With pigeons, the element A does eventually gain control over responding and a simultaneous discrimination occurs within the AB- display. The pigeons no longer respond to B but fail to withhold responding to A on the AB- trials. The P(A) on S- trials remains approximately equal to the P(A) on S+ trials (Jenkins, 1973). One explanation is that responding persists to the S- display because it is composed of elements common the S+ display which are controlling responsing (Jenkins & Sainsbury, 1970). The A responses are thus maintained by partial reinforcement.

The phenomena of locating responses to a particular element even though the location of the element is switched from trial to trial is called feature tracking. In the FP case the organisms come to locate responses on the B feature and track the feature from quadrant to quadrant and from side to side on the stimulus display. Jenkins and Sainsbury (1969, 1970) explain that the organism responds to the stimulus with the greatest probability of reinforcement. Jenkins and Sainsbury (1969) made an assumption that the effects of reinforcement were restricted to the stimulus element to which the response was directed on any trial. This assumption was made to explain the eventual shift from A to B within the AB+ display. Jenkins (1974) states that such an assumption is suspect. He found that B could still accumulate excitatory strength even when B responses were not reinforced. Therefore, to assume that A does not gain excitatory strength when responses are directed to B may be unwarranted. Jenkins and Sainsbury (1969) proposed a theory of complete selectivity. Now, Rescorla and Wagner (1972) have a theory that incorporates an opposite assumption.

The Rescorla-Wagner (1972) model was developed as a theory for Pavlovian conditioning that can be applied to instrumental conditioning. In this model the associative strength of a compound stimulus AB depends not only upon the current total strength of A and B but on the strength of one component relative to the other. When an AB compound is followed by an unconditioned stimulus, the excitatory value of B will be increased more if A is arranged to have low excitatory value than when A has high excitatory value. Also, when AB trials are nonreinforced, the conditioned response inhibition characteristics of B will increase much more if A has high excitatory value than when A has low excitatory value. The theory is represented in the following mathematical model:

$$\Delta V_A = \alpha A \ \beta 1 \ (\lambda 1 - V_{AB})$$

$$\Delta V_B = \alpha B \ \beta 1 \ (\lambda 1 - V_{AB})$$

$$V_{AB} = V_A + V_B$$

V = associative strength

β = learning rate parameter associated with the reinforcer

λ = the asymptotic level of associative strength which the reinforcer will support

α = the learning rate parameter related to the salience of the component

Wagner and Rescorla assume that the effects of reinforcement or extinction on the excitatory value of either element of a compound stimulus is a function of the total excitatory value of the compound but is entirely independent of the relative excitatory values of the components.

Although the Wagner-Rescorla model does not deal with feature tracking (location of responses on the B element) it does predict the major results of the Jenkins (1973) study. Jenkins manipulated the consequences of B responses and cue salience and found three factors that influence responses to the B element. One factor is whether or not pecks directed to B are reinforced. A second and most important factor is the role of B as a signal that A responses will or will not be reinforced when B responses are not reinforced. A third factor is the salience of the B stimulus relative to the A stimulus. The more salient B the greater the probability that responses will be made to it. Jenkins used six groups. The groups are arranged in Table 1 according to the expected number of _B_ responses based on the cue function of _B_ and the probability of a response to _B_ being reinforced. The arrangement represents a continuum with the most _B_ responses emitted in the feature-positive condition to the least _B_ responses occurring in the feature-negative condition. The Wagner-Rescorla model would predict the same ordering of the groups.

According to Wagner-Rescorla model both _A_ and _B_ will gain strength in the FP condition. This assumption that reinforcement strength is not confined solely to the element responded to resembles a nonselection theory. However, _B_ will gain strength on every trial while _A_ will gain strength on _A_+_B_+ trials but will lose strength on A- trials. The element _A_ does not recover enough strength on _A_+_B_+ trials to overcome the strength lost on _A_- trials. The result is that the strength of _A_ approaches zero as _B_ gains in strength. In the _A_+_B_+/_A_- condition, _A_+ responses cause both _A_ and _B_ to gain in strength while _A_- responses cause _A_ to lose strength. The _B_ responses cause a loss in strength for both _A_ and _B_. The _A_ and _B_ elements both gain and lose excitatory strength on each trial, but since only responses directed to _A_ are reinforced _A_ eventually gains more strength than _B_. The lower excitatory value of _B_ in conditions _A_+_B_- and _A_+_B_-/_A_+ is correctly predicted by

Table 1

Stimuli Displayed and the Outcome of Responses During Training for Each Group in Jenkin's 1974 Study

Group	Comment
1) A(+)B(+)/A(-)	The feature-positive case. B signals that response to A is reinforced and the B-response is also reinforced.
2) A(+)B(-)/A(-)	B signals that response to A is reinforced, but the B-response is nonreinforced.
3) A(+)B(-)	B is present when response to A is reinforced, but the B-response is nonreinforced.
4) A(+)B(-)/A(+)	B is sometimes present when a response to A is reinforced, but the B-response is nonreinforced.
5) A(+)/A(-)	B not present. Used to test generalization from A to B on test.
6) A(+)/A(-)B(-)	The feature-negative case. B signals that response to A is nonreinforced, and the B-response is nonreinforced.

Wagner-Rescorla model. Here, A gains strength on every trial and approaches asymptote strength while B's strength remains low. In the feature-negative condition, A+/A-B- condition, the theory predicts that B will accumulate inhibitory strength. Theoretically, the strength of the A-B- compound should approach an asymptote strength of zero due to non-reinforcement. For this to happen B must acquire inhibitory strength to overcome the excitatory strength that A gets from A+ trials. None of the generalization tests indicated that B gained inhibitory strength. This study also yielded results that do not support the Wagner-Rescorla model of non-selection. In the A+B-/A condition, significantly fewer responses were made to the less salient green dot than to the more salient red dot. Apparently B responses are reduced more by non-reinforcement of B responses on the A+B- display than by non-reinforcement of a response to A on that display. As Jenkins points out, the shift to B was prevented by a very small number of nonreinforced B responses when B was the less salient dot and it is most unlikely that the same number of nonreinforcements delivered without regard to which stimulus was pecked would have prevented the shift to B.

Jenkins (1973) conducted an experiment to demonstrate that the element responded to is more affected by reinforcement or nonreinforcement than the element which is not selected. Jenkins set up a simultaneous discrimination between the elements on an A+B- display. The B element was more salient and received an initial high rate of responding but a preference for A gradually developed. In this demonstration by Jenkins, reinforcement appeared to selectively increase excitatory values and nonreinforcement seems to selectively decrease excitatory values of elements to which a response is made.

Hull's (1943, 1952) theory was one of reinforcement selection. The strength of a component was thought to be an independent value unaffected by the strength of the other components. Rescorla and Wagner (1972) assert that strength of one stimulus component depends upon the total associative strength of the compound of which it is a part. It appears that neither of the two theories in their present form can adequately account for all the data.

Noncontinuity theory is the second type of discrimination theory. It developed out of the use of compound conditioned stimuli in the classical conditioning paradigm. Compound stimuli have assumed an important role in the investigation of operant discrimination learning. A compound stimulus is a combination of independent stimulus elements, cues, or features (i.e., light + tone; red square). The cues of a compound stimulus can be relevant (correlated with reinforcement) or irrelevant (not correlated with reinforcement). If both cues of a compounds stimulus are relevant then the two cues are redundant. "If a subject has been trained with several redundant relevant cues serving for differential behavior, later tests may show that behavior is primarily under the control of only one component of the entire stimulus complex (Trabasso & Bower, 1968, p. 527)." They contend that organisms do not attend equally to all stimulus dimensions in solving a discrimination.

If organisms are reinforced for responding to a redundant relevant stimulus complex _AB_, some of the organisms apparently base their discrimination solely on cue _A_, others solely on cue _B_, and still others on both cues _A_ and _B_ (Suchman & Trabasso, 1966; Sutherland & Mackintosh, 1964). In some cases, the effect that one element has on responding can overshadow any effect that the other might have had if it had been conditioned singly. For example, if a pigeon is conditioned to peck a key in the presence of a light + tone, the pigeon's responding may be significantly affected by modification or removal of the light but not by modification or removal of the tone. Broadbent (1958) proposes that some dimensions of a complex stimulus gain control over a response whereas others do not because organisms are limited in the number of stimulus properties to which they can attend simultaneously. This is the concept of limited-channel capacity. When an organism's behavior is controlled by only certain dimensions of a stimulus complex, selective attention is the phenomena said to be occuring.

Thus, there are two basic assumptions in attention theory. The first assumption is that organisms will learn only about those cues to which they are attending. The second assumption is that organisms are limited in the number of stimuli to which they can attend simultaneously. Early noncontinuity theorists such as Lashley (1942) claimed that only one cue could be attended to at a time. Recently two-stage models of discrimination learning have developed. The two-stage model states that discrimination learning is a chained response. The organism must first learn to attend to the relevant stimulus dimension and then secondly the organism must learn to attach an appropriate response to the correct cue. Since some elements are common to both the S+ and S- displays a neutralization of these common elements must occur if the discrimination problem is to be solved. Two-stage theories of discrimination learning attempt to cope with this problem of overlap. These theories might be grouped into three categories.

One type of two-stage theory was developed by Wyckoff (1952). He introduced the observing-response concept. An observing response is any response which produces exposure to a discriminative stimulus. This conception of attention is more ameliorative to a S-R paradigm. The proposed orienting responses precedes the operant response in a behavior chain. The first link is the response exposing the subject to the discriminative stimulus; the second link is the response to the discriminate stimulus. If the second link is reinforced it establishes itself as a conditioned or secondary reinforcer. The observing response is thus strengthened or weakened by the consequences of the last link in the behavior chain. If reinforcement follows the last link, the observing response is likely to be repeated. If nonreinforcement follows the operant response the observing response will be weakened.

A second type of two-state attention theory is the mediating response model (Kendler & Kendler, 1962). However, there remains some confusion as to what a mediating response is. Some authors such as Hill (1971) state that the observing response is a mediating response. However there are some instances in which attention cannot be characterized by an overt orienting response. Some animals attend without obvious overt orientation responses. This situation may have led some theorists to adopt a mediating response position. Still, the mediating response is not well defined. Whatever its nature the mediating response can be viewed as a link connecting the external stimulus and the observable response. For this model discrimination learning is still represented as a behavior chain.

The third type of model is the attention response model. Goodwin and Lawrence (1955) suggest that the subject must learn to attend to the relevant dimension. Sutherland (1959) states that organisms must learn to switch in the relevant analyzer. Lovejoy (1966) and Zeaman and House (1963) state that the subject must learn attention responses. Lawrence (1949, 1963) interprets discrimination learning in terms of acquired distinctiveness of cues or stimulus coding. Some of these theories sound very cognitive in nature and suggest that the subject tries out a number of hypotheses and then discards or reuses them according to the successfulness. There is not a clear line of demarcation between the types of two-stage attention theories.

The attention theories make an assumption of selection. The effects of reinforcement and nonreinforcement are hypothesized to be confined to only the stimulus dimension that is attended to on any one trial. Sutherland (1959) stated that failure of a dimension to provide accurate prediction of outcome would cause the organism to shift attention to another feature. Restle (1962) states that resampling of dimensions is the result of nonreinforcement. Lovejoy (1966) proposes that attention to an element is strengthened on reinforced trials and weakened on nonreinforced trials. The assumption of selection makes attention theory similar to Jenkins and Sainsbury's (1969) original position. Jenkins (1973) has modified his position but still accounts for discrimination learning without using an attention response concept.

Jenkins and Sainsbury (1969) compared the cue acquisition phase of the two-stage attention theory to what they termed a "search theory". This theory states that in a successive discrimination, non-reinforcement induces search in the form of an elemental response shift. A response to an element that is not reinforced is weakened and a new response is more likely to occur at the next trial. In the feature-positive condition, search will strengthen both search and the response to B. In the feature-negative-condition, reinforcement following search strengthens search and the response to A. An increase in the tendency to respond to A will compete with the tendency to search on the next trial. Search following nonreinforced A-responses would yield B which has never been reinforced. Nonreinforcement is less damaging in the feature-positive case.

In the feature-negative condition, non-reinforcement for search occurs only on trials in which the response is made to B. Reinforcement increases the probability of search more in the feature-positive than in the feature-negative condition and non-reinforcement decreases the probability of search more in the feature-negative than in the feature-positive case. It is hypothesized that search will be maintained throughout training in the feature-positive condition but will extinguish in the feature-negative condition.

If a failure to develop differential performance in the feature-negative condition were caused by a failure to search for the distinctive feature, one would expect no discrimination within the negative display between common and distinctive elements. Jenkins and Sainsbury (1969) and Jenkins (1973) found that their feature-negative subjects eventually located all responses to the common feature. A search theory stated that a successive discrimination fails to develop in the feature-negative condition because the outcome of the search phase that is required to make the B-feature distinctive from the rest of the display is not supported. However, the failure of a successive discrimination to develop in the feature-negative condition can better be accounted for by the subject's inability to form a conditional discrimination rather than failure to develop an observing response that alters the availability of the features on the display.

This study will investigate the effects of the proximity of elements upon the feature-positive effect with preschool children. Sainsbury (1971a) varied the proximity of colored dots in a feature discrimination paradigm with pigeons and found that when the elements were compacted into one quadrant rather than centered in each of the quadrants, the feature-positive effect was significantly reduced. In Sainsbury's (1971a) study the stimulus displays were divided into four equal parts by a 1/16 inch metal strip. In the distributed condition the elements were located in the centers of each quadrant. In the compacted condition all the elements were located in one quadrant. This stimulus arrangement changed more than proximity because the elements were no longer divided by the metal strips and feature touching could not be recorded.

This experiment will attempt to improve upon Sainsbury's (1971a) design. Instead of compacting the elements into one quadrant, the elements will be compacted toward the center of the display while still being divided into quadrants by two intersecting lines. This way feature touching can still be recorded. A second purpose of this study will be to investigate cue-response separation in the feature discrimination paradigm. One successful strategy that a subject can use to solve the feature-negative discrimination is to locate the distinctive feature and then respond to the other display. Sainsbury (1973) reported several subjects apparently employing this strategy. It is hypothesized that the difficulty of acquiring this response prevents many subjects from successfully using this strategy. Previous studies have shown that manipulating the spatial separation of cue, response and reward can affect the acquisition of a discrimination in primates (c.f., Murphy & Miller, 1955, 1958, 1959; Stollnitz & Schrier, 1962). Jeffrey and Cohen (1964) found this effect with children but Campione and Beaton (1972) did not. It is hypothesized that primates restrict their attention to where they put their fingers (Schuck, 1960).

It is further hypothesized that the elements touched indicate the cues used to solve the discrimination. In the feature-positive case an organism can solve the discrimination by using B as a cue to respond. Cue and response can be separated in the S+ display by reinforcing only A-responses on the AB display. Learning rate appears to be a direct function of the distance between cue and response. The greater the separation of cue and response the greater the difficulty in acquiring the discrimination. This effect was found with rhesus monkeys (c.f., McClearn & Harlow, 1954; Meyer, Polidora & McConnell, 1961; Miller & Murphy, 1964; Murphy & Miller, 1959).

A third purpose of this study will be to investigate the stimulus identification hypothesis. This hypothesis states that the organism will solve the discrimination either by grouping or not grouping the elements and that this strategy is relatively fixed. In other words, the subject can base his discrimination on the individual elements or on a pattern of elements. A generalization test will be given all subjects to investigate the dimension(s) of the stimulus complex that controlled responding during the experiment.

CHAPTER II

METHOD

Subjects

The subjects were sixty 3-5 year-old children from preschools and day care centers in Des Moines, Iowa. The subjects attended either Wakonda Christian Church Preschool, Calvary Christian Church Day Care Center, or Children's Garden.

Apparatus

Two stimulus cards were presented to the subjects on an 18" long by 10" high metal card holder. The holder was black with two angle supports on each side so that the stimulus cards could be loaded on the holder, out of view of the subject, and then flipped over to expose both cards simultaneously. The holder tilted at about a 120 degree angle away from the subject. Eight mirror mounting brackets held the stimulus cards in position. The brackets were positioned such that the center to center distance of the stimulus cards was 10.5 inches.

Stimuli and Recording

The stimulus cards were 6" by 6" white cardboard. Each card had two perpendicular black lines intersecting in the middle of the card which divided it into four 3" by 3" quadrants. The two pretraining stimulus cards (see Figure 2)

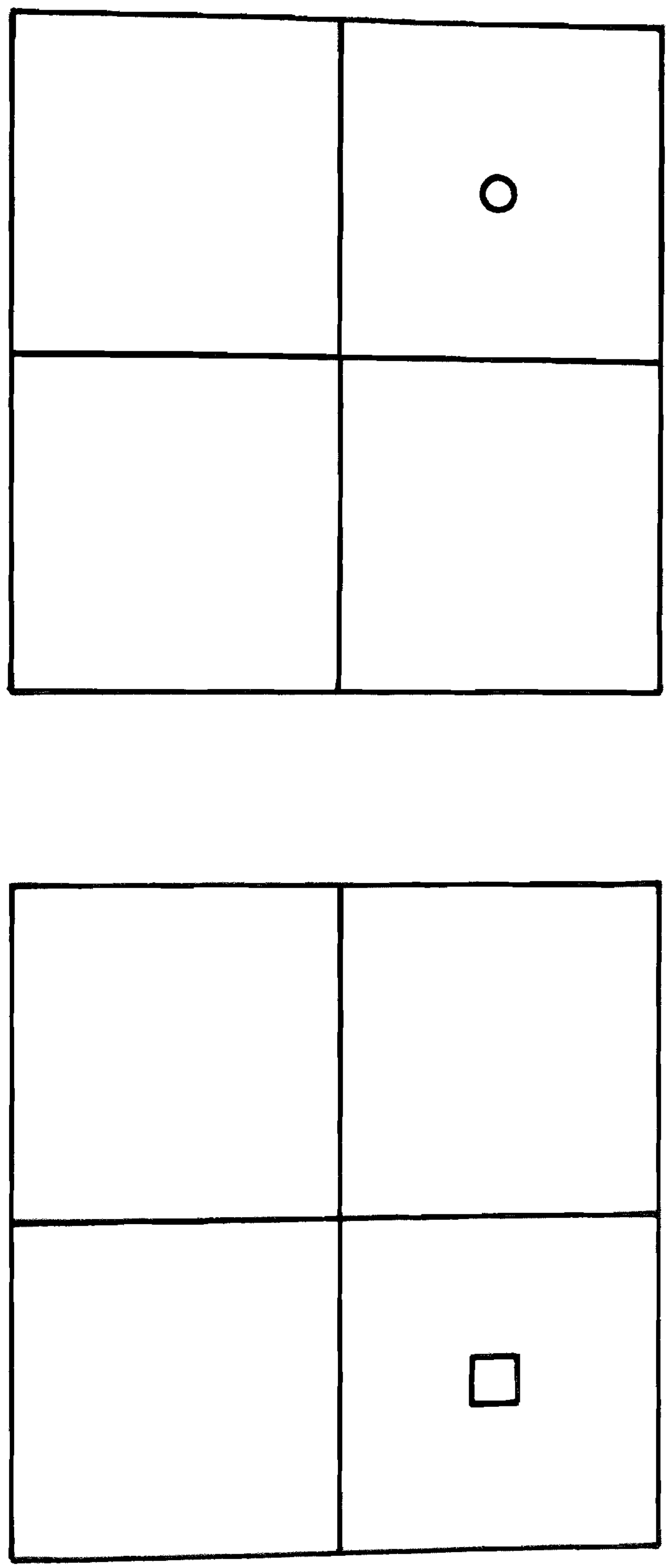

Fig. 2. Pretraining stimuli.

each had one 3/8" geometric figure centered in one quadrant. The two figures used were a circle with a 3/8" diameter and a square with 3/8" sides. The sizes of these figures remained the same throughout the experiment. After pretraining, subjects received differential training between a common feature card and a distinctive feature card (see Figure 3). The common feature card had four identical geometric figures. The distinctive feature card continued 3 figures exactly like those on the common feature card plus one different geometric figure. On either card each of the four elements occupied a separate quadrant. Half of all subjects were trained with a square as the distinctive feature and half with the circle as the distinctive feature. In each of these cases, half of the subjects were trained with cards which had the elements located in the outer corners of each card and half with the elements located near the centers of the cards. The distance from the middle of a card to the middle of any geometric figure was 3.84 inches for the distributed elements and .37 inches for the compacted elements.

In the generalization test all of the elements were centered in the quadrants (see Figure 4). The distance from the center of the figure to the center of the card was 2.12 inches. In the new elements generalization test, the elements used were 3/8 inch triangles and hexagons.

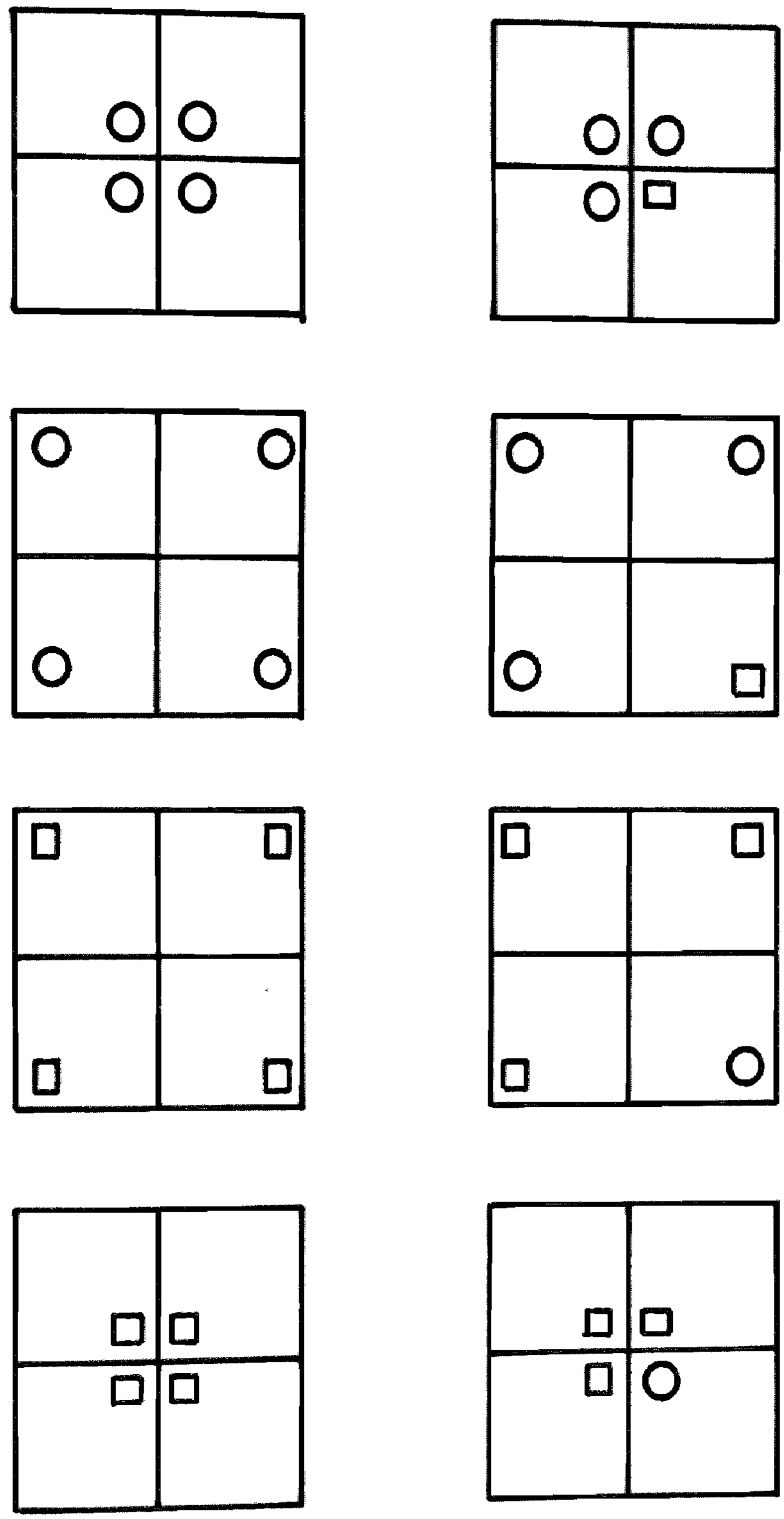

Fig. 3. Training Stimuli.

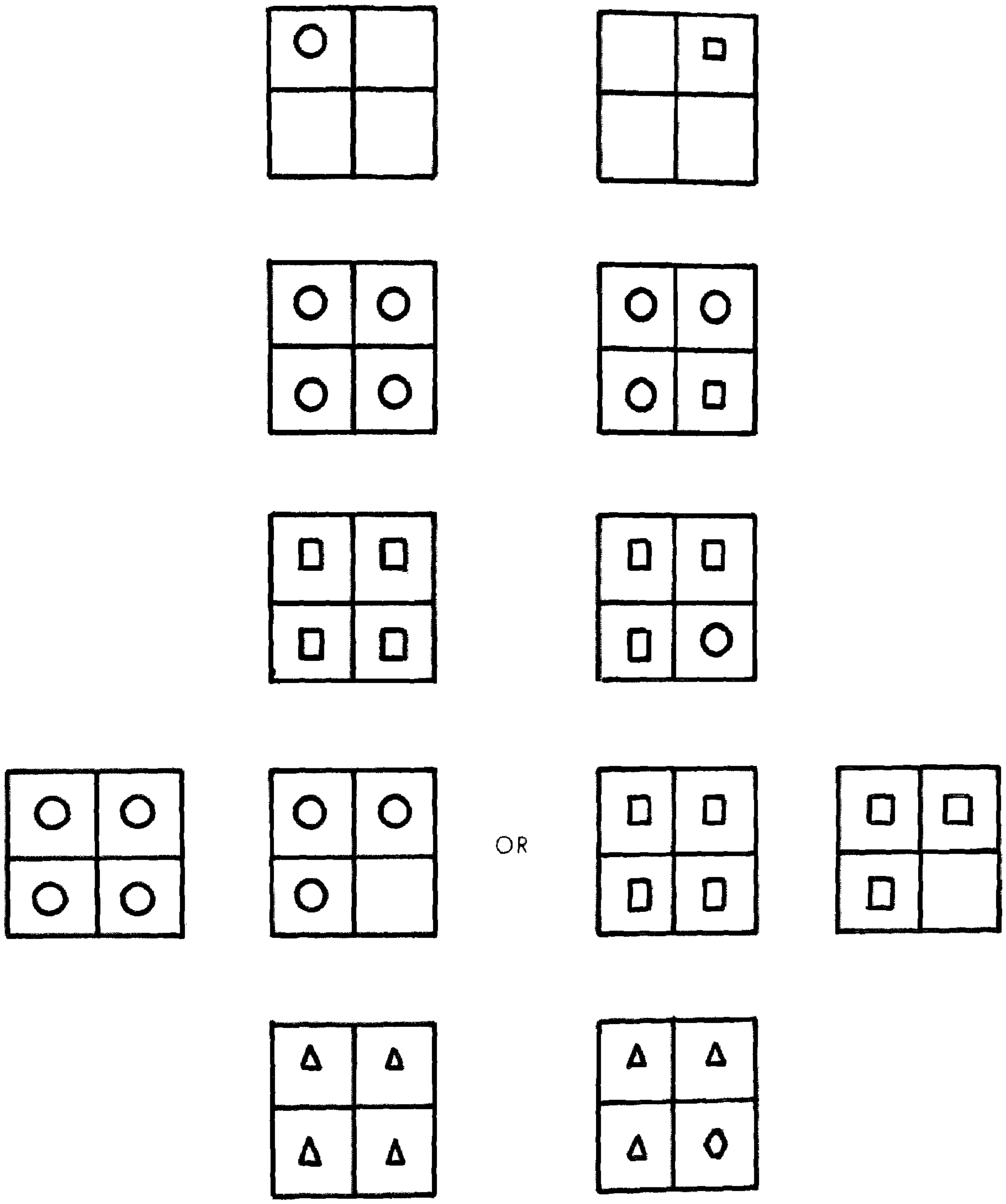

Fig. 4. Generalization test stimulus displays.

A trial was begun with the presentation of the stimulus display and terminated when the subject touched one of the elements or when ten seconds elapsed--whichever came first. The location and the type of element touched was recorded for each trial.

Except during the generalization test, verbal feedback was given for each response. When the subject made a correct response the experimenter would say, "Yes, that is the one" and then remove the display. When the subject made an incorrect response the experimenter would say, "No, that is not the one" and then remove the display. When each session was completed, the experimenter thanked the subject for playing the game and the subject was shown a number of small toys from which he or she could choose one to keep.

Procedure

The subjects were randomly assigned to one of twelve experimental conditions. Each subject was brought to an area of their preschool or day care center that was quiet and where other children would not interfere with the procedure. The subject was seated at a table directly across from the experimenter and facing him. Each subject was given the following instructions at the beginning of the first session:

> "We are going to play a guessing game. I am going to show you two cards (experimenter presents cards). The cards have shapes on them.

I want you to touch the shape that you think is the good shape. Touch it with your finger."

Discrimination training then started. Pretraining consisted of ten trials or five consecutive correct responses. Table 2 lists the twelve groups and the type of pretraining each received. The purpose of pretraining was to insure that the subject touched the elements and to reinforce the subject for touching the same type of element necessary to solve the training phase discrimination.

In the training phase, subjects in the feature-positive-one conditions (FP-1) were reinforced for touching any element (geometric figure) on the display with the distinctive feature while subjects in the feature-positive-two condition (FP-2) were reinforced for touching any common element on the distinctive-feature display and subjects in the feature-negative condition (FN) were reinforced for touching any of the elements on the common feature display. The stimulus cards were presented according to a preset randomized sequence. Over forty trials, each card appeared on either side of the display half of the time. Also, the distinctive feature appeared in each of the quadrants equally often. Between trials the experimenter recorded responses and changed the stimulus display. This resulted in an intertrial interval of approximately 10 to 15 seconds.

In a later session, which was conducted the following day, the subjects were given a generalization test. The

Table 2

Pretraining Phase S+ Elements
for Each Group

	TRAINING PHASE B-FEATURE	FEATURE POSITIVE I	FEATURE POSITIVE II	FEATURE NEGATIVE
DISTRIBUTED ELEMENTS	CIRCLE	CIRCLE	SQUARE	SQUARE
	SQUARE	SQUARE	CIRCLE	CIRCLE
COMPACTED ELEMENTS	CIRCLE	CIRCLE	SQUARE	SQUARE
	SQUARE	SQUARE	CIRCLE	CIRCLE

instructions to the subjects were:

> "Do you remember the guessing game that we played the last time? Well, we are going to play the same game and we will use some new cards."

In this session the experimenter briefly reviewed the discrimination with the subject. This review consisted of no more than ten trials. Positive and negative feedback was given during the review trials. The experimenter then presented five stimulus arrangements (see Figure 4) eight times in random order for forty trials. The generalization tests consisted of the following stimulus arrangements:

1. The original pretraining stimulus cards.
2. The training phase stimulus arrangement.
3. A reversal of the training stimulus arrangement (common feature now distinctive).
4. A feature discrimination with novel elements (A = triangles, B = hexagon).
5. A three versus four common elements arrangement.

CHAPTER III

RESULTS

Three measures of performance were taken during the training phase of the experiment. One measure was the total number of errors for each child. A second measure was the percentage of distinctive feature responses (_B_) on the distinctive feature display (_AB_) for trials 1-10. This measure is represented as the ratio _B/AB_. A third measure was the percentage of total responses made to the distinctive feature (_B_). This measure is represented as the ratio _B/A+B_. Each statistic in this analysis was computed by dividing the total number of responses to the distinctive feature by the total number of "actual" training trials (not always 40). The first performance measure is a measure of error and the other two are measures to investigate feature tracking.

The training phase of the experiment was terminated after 10 consecutive correct responses or 40 trials--whichever came first. Those children who reached criteria in less than 40 trials were assumed to perform at the 100% correct response level for the remaining trials. The actual number of training trials received for all children ranged from 10 (the least possible) to 40 (the maximum possible). The 40 trials were divided into 4 blocks of 10 trials each. Errors were recorded as the total number occuring in each block (see Appendices A, B, C).

An analysis of variance of errors (see Table 3) resulted in no significant difference in the groups trained with either the circle or the square as the distinctive feature and therefore these two conditions will be combined for analysis. A significant difference among the feature conditions was obtained, (F (2.48) = 9.9019, p < .01). Tukey's HSD test results (see Table 4) revealed that all feature groups differed significantly from each other in the distributed elements condition but only the FP-2 and FN groups differed significantly in the compacted elements condition. Thus, a feature-positive effect was obtained only for the distributed elements condition.

The children in the compacted elements condition made fewer errors than the children in the distributed elements condition, (F (1.48) = 4.7440, p < .05). Tukey's HSD test (see Table 4) revealed that compacting the elements significantly reduced errors for both the FP-2 and FN conditions but not for the FP-1 condition.

A graph of errors (see Figure 5) shows that the FP-2 group made the most errors in both distributed and compacted elements conditions. The FN group made more errors than the FP-1 group in the distributed elements condition but fewer in the compacted elements condition. However, a feature x proximity interaction was not significant.

In the within source of variance there was a

Table 3

Analysis of Variance of Error Responses

Source	SS	df	MS	F
Between	1,254.1500	59		
Features (F)	282.1000	2	141.0500	9.9019**
Proximity (C)	68.2667	1	68.2667	4.7440*
Elements (D)	.0667	1	.0667	.0046
F X C	84.9333	2	42.4666	2.9511
F X D	9.7333	2	4.8666	.3382
D X C	43.3499	1	43.3499	3.0125
F X D X C	75.1004	2	37.5502	2.6095
Between Error	690.6000	48	14.3900	
Within	484.5000	180		
Blocks (B)	135.6500	3	45.2167	25.7763**
F X B	49.0000	6	8.1667	4.6555**
D X B	6.7333	3	2.2444	1.2794
C X B	.4000	3	.1333	.0760
D X C X B	4.1834	3	1.3945	.7949
F X C X B	7.5000	6	1.2500	.7126
F X D X B	15.1667	6	2.5278	1.6393
F X D X C X B	13.2666	6	2.2111	2.2605
Within Error	252.6000	144	1.7542	
Total	1,738.6500	239		

* $p < .05$
** $p < .01$

Table 4

Tukey's HSD Test

	$\bar{X}_6$	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_5$	$\bar{X}_4$	$\bar{X}_3$
$\bar{X}_6 = 2.4$		1.1	3.5	6.8**	7.8**	16.2**
$\bar{X}_1 = 3.5$			2.4	5.7*	6.7**	15.1**
$\bar{X}_2 = 5.9$				3.3	4.3	12.7**
$\bar{X}_5 = 9.2$					1.0	9.4**
$\bar{X}_4 = 10.2$						8.4**
$\bar{X}_3 = 18.6$						

* $p < .05$
** $p < .01$

$\bar{X}_1$ = FP-1 Distributed Elements

$\bar{X}_2$ = FP-1 Compacted Elements

$\bar{X}_3$ = FP-2 Distributed Elements

$\bar{X}_4$ = FP-2 Compacted Elements

$\bar{X}_5$ = FN Distributed Elements

$\bar{X}_6$ = FN Compacted Elements

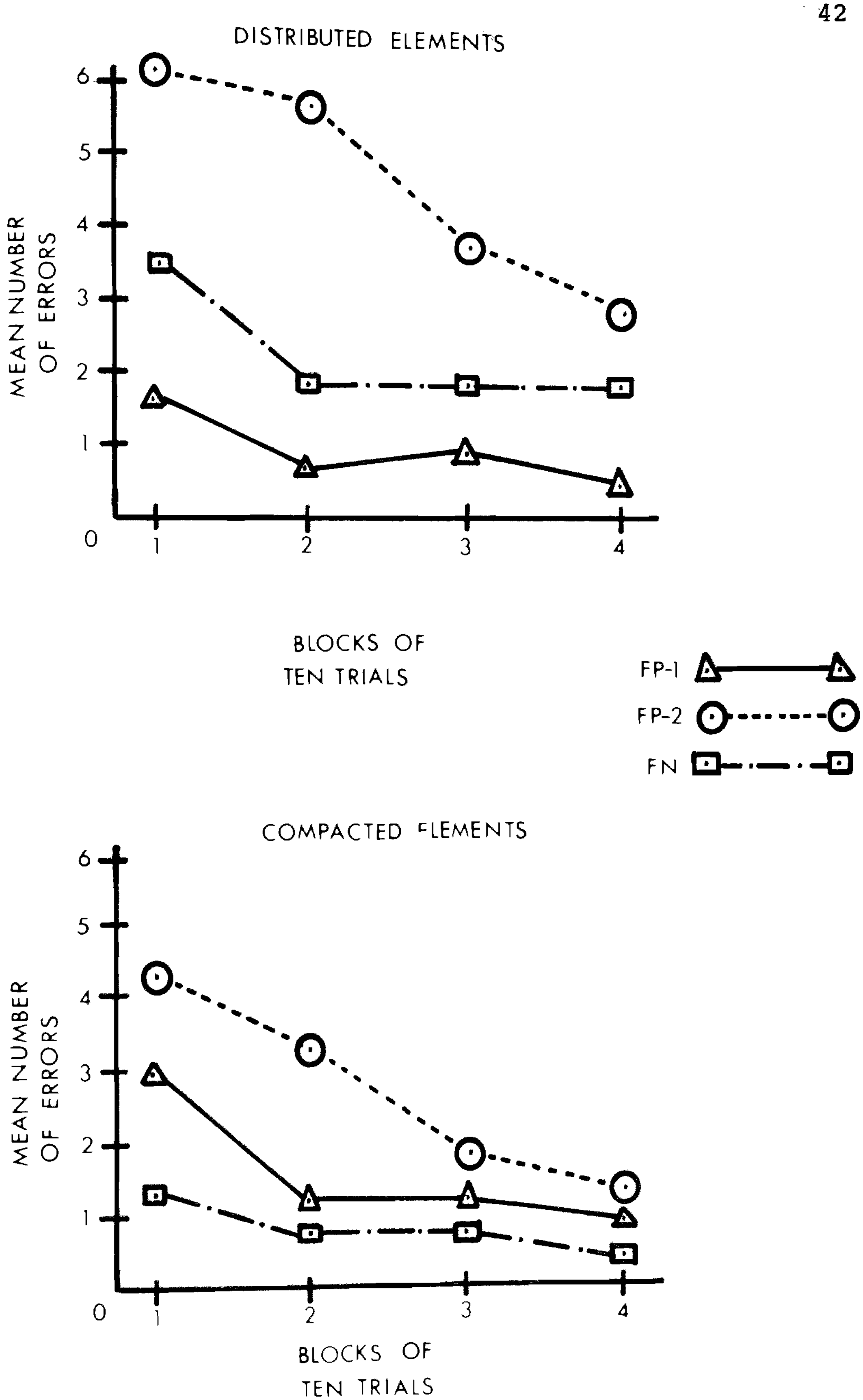
DISTRIBUTED ELEMENTS
MEAN NUMBER OF ERRORS
6
5
4
3
2
1
0
1
2
3
4
BLOCKS OF TEN TRIALS
FP-1
FP-2
FN
COMPACTED ELEMENTS
MEAN NUMBER OF ERRORS
6
5
4
3
2
1
0
1
2
3
4
BLOCKS OF TEN TRIALS

Fig. 5. Errors.

significant difference within blocks of trials, (*F* (3.144) = 25.7763, *p* < .01). Also a feature x blocks interaction was significant, (*F* (6.144) = 21.3716, *p* < .01). The children made fewer errors across trials depending upon the condition to which they were assigned. No other significance was found.

An analysis of variance of the percentage of *B* responses on the *AB* card for trials 1-10 (see Table 5) yielded a significant difference only among the feature conditions, (*F* (2.48) = 21.3716, *p* < .01). Figure 6 is a bar graph of mean percentage of *B*/*AB* responses for trials 1-10. The graph shows that *B*-responses were highest in the FP-1 condition where *B*-responses were reinforced. The FP-1 group differed significantly from the FP-2 and FN groups (Tukey's HSD test).

The mean percentage of *B*-responses for total training trials (*B*/*A*+*B*) was also analyzed (see Table 6). The only significance was among the feature conditions, (*F* (2.48) = 36.6918, *p* < .01). Figure 7, a bar graph of the mean percentage of *B*-responses to total responses, shows that *B*-responses were highest in the FP-1 condition where responses to *B* were reinforced. A Tukey's HSD test showed that the FP-1 group emitted significantly more *B*-responses than either the FP-2 or FN groups.

The results of the generalization tests are recorded in Appendices D, E, and F. Responses to the original "OTS",

Table 5

Analysis of Variance of the Ratio of B/AB
Responses for Trials 1-10

Source	df	SS	MS	F
Feature (F)	2	3.7613	1.8807	21.3716**
Elements (D)	1	.0055	.0055	.0625
Proximity (C)	1	.2707	.2707	3.0761
F X D	2	.0297	.0149	.1693
F X C	2	.0743	.0372	.4227
D X C	1	.3067	.3067	3.4852
F X D X C	2	.0941	.0471	.5352
Error	48	4.2248	.0880	
Total	59	8.7671		

* $P < .05$
** $P < .01$

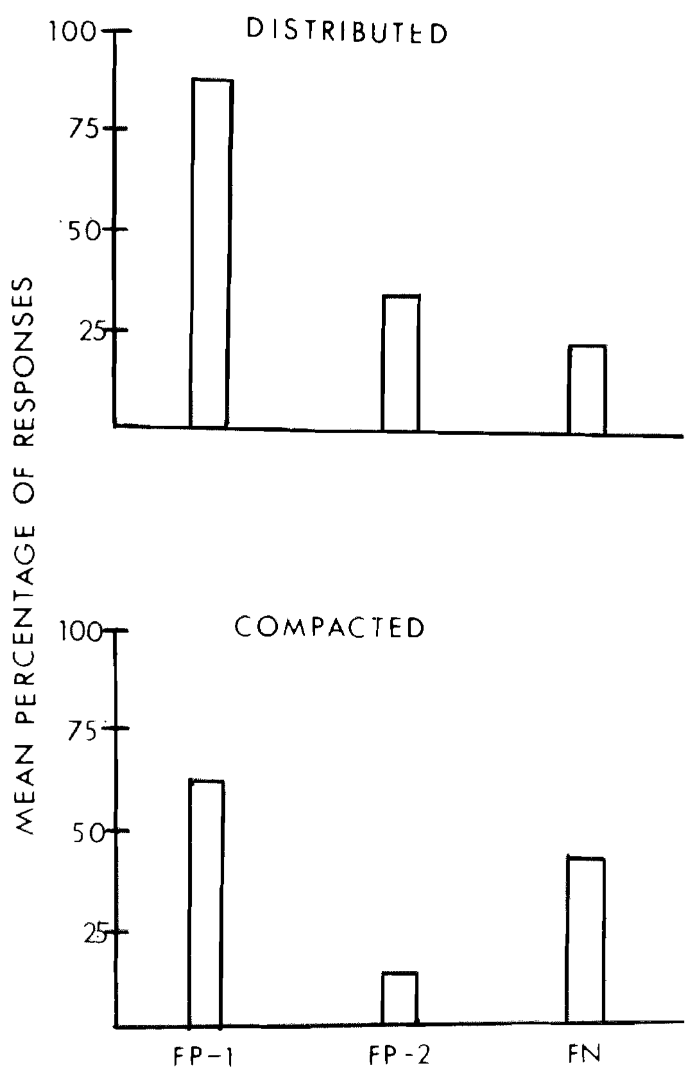

Fig. 6. Bar graph of mean percentage of B/AB responses for trials 1-10.

Table 6

Analysis of Variance of the Ratio of B/A Responses During the Training Phase

Source	df	SS	MS	F
Feature (F)	2	3.8821	1.9410	36.6918**
Element (C)	1	.0034	.0034	.0642
Proximity (D)	1	.1852	.1852	3.5009
F X D	2	.0734	.0367	.6937
F X C	2	.2549	.1274	2.4083
D X C	1	.1109	.1109	2.0964
F X D X C	2	.0117	.0058	.1096
Error	48	2.5412	.0529	
Total	59	7.0628		

* $p < .05$
** $p < .01$

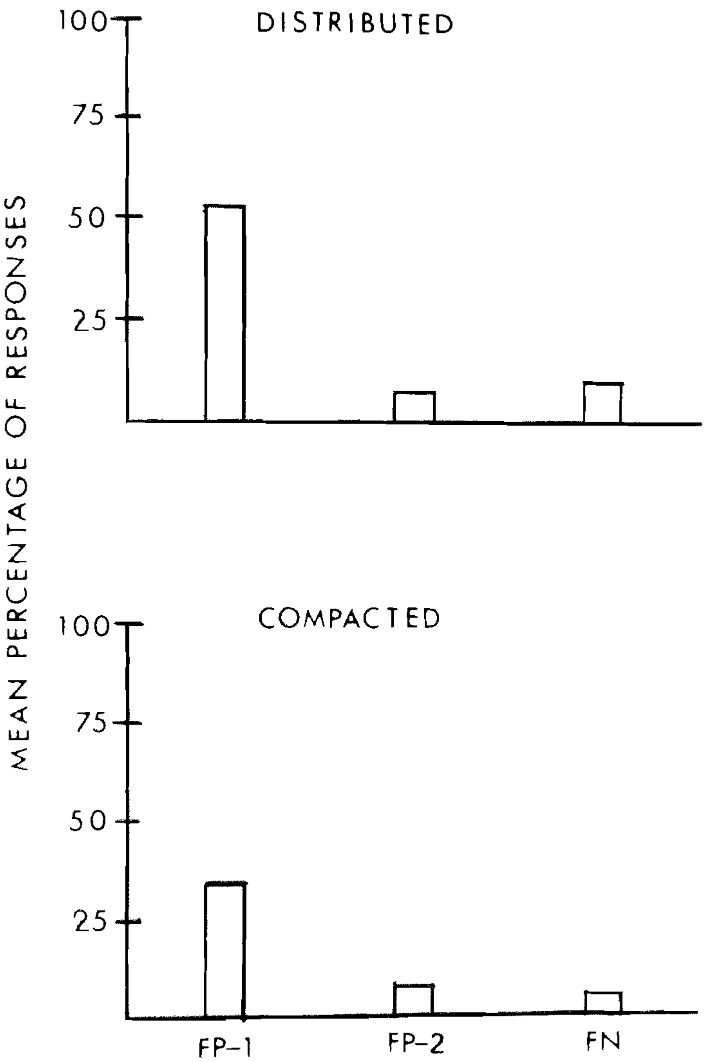

Fig. 7. Bar graph of mean percentage of B-responses for total training trials.

reversed "REV", and new "NEW", elements generalization tests are represented by the schema AB/A. The letter B represents the distinctive feature, the letter AB represent the distinctive feature card, and the letter A represents the common feature card.

Figure 8 is a composite of bar graphs for the original, reversed, and new elements generalization tests. All groups responded as expected to the original elements generalization tests. Feature-positive children showed a preference for the AB card while the feature-negative children preferred the A-card. In the reversed elements test (REV) the feature-positive children continue to respond to the AB display. In the feature-negative condition, the children switched preferences from the (OTS) test and responded more to the AB display. This switch is most obvious in the distributed elements condition. In the new elements (NEW) generalization test the feature-positive children continue to respond to the AB display while the feature-negative children respond more to A alone if the display elements are compacted.

Figure 9 is a bar graph of the "OTS", "REV", and "NEW" elements generalization tests results combined for analysis. It appears that overall the feature-positive children chose the AB card and the feature-negative children chose the A card when the elements were compacted but showed no preference when the elements were distributed.

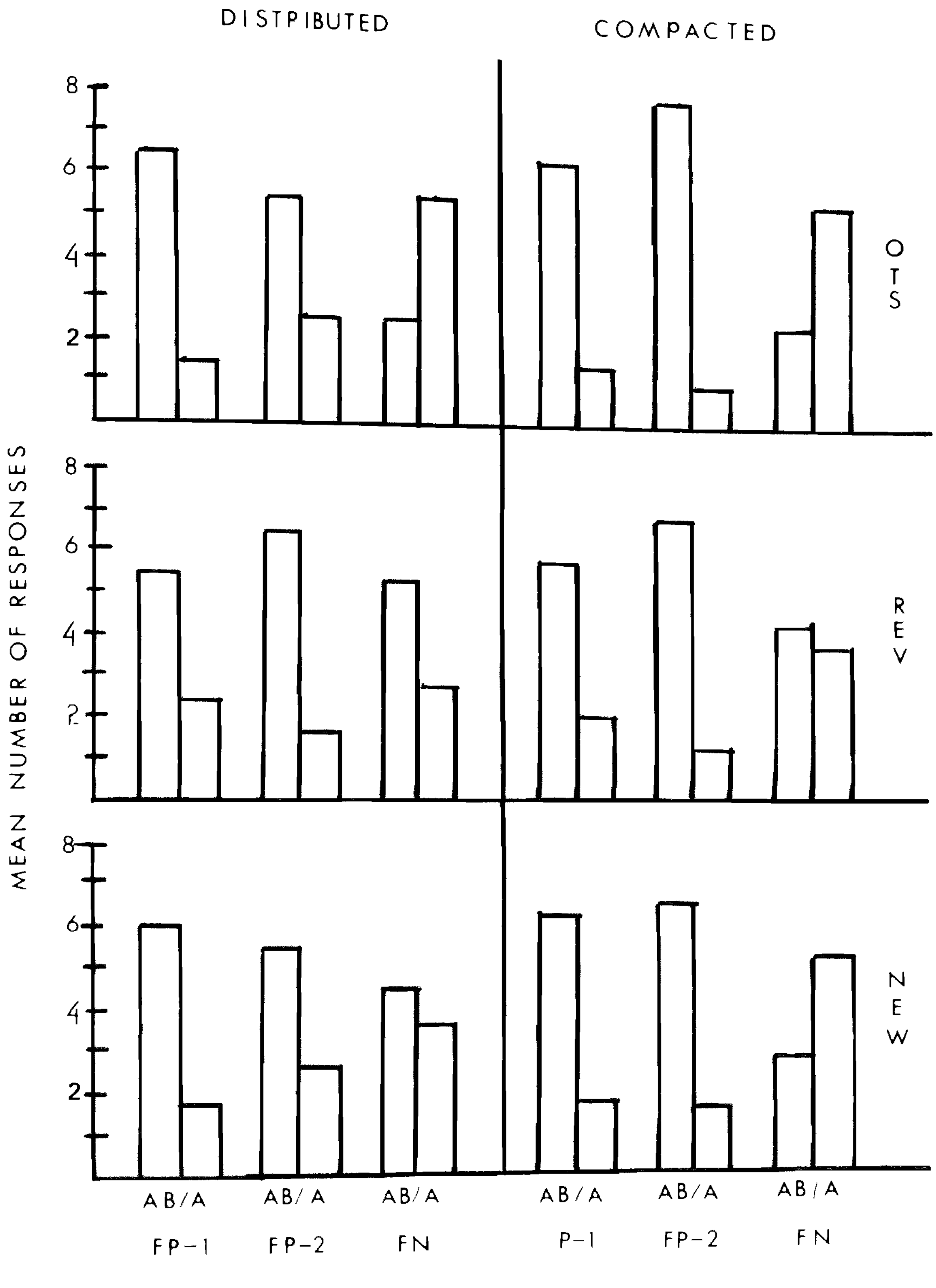

Fig. 8. Bar graph of original elements (OTS), reversed (REV), and new (NEW) elements generalization test results.

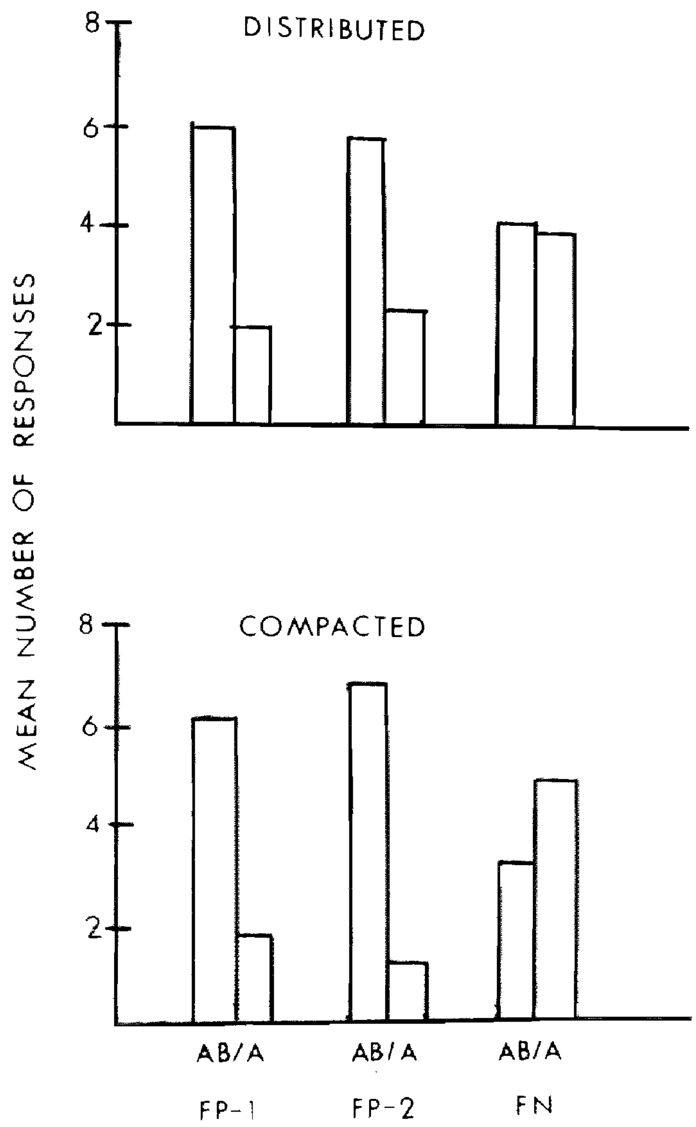

Fig. 9. Bar graph of the original, reversed, and new elements generalization tests results combined for analysis.

Figure 10 is a bar graph of the original pretraining stimulus test results. It appears that all children maintained a preference for the pretraining S+ stimulus.

Figure 11 is a bar graph of the three versus four common elements generalization test. The feature-positive children chose the three common elements card while the feature-negative children chose the four common elements card. Only two children responded to the blank sector on the three common elements card. Both children were in the FP-1 condition. One child in the distributed elements condition responded exclusively to the blank sector, while the other child in the compacted elements condition emitted only one response to the blank sector.

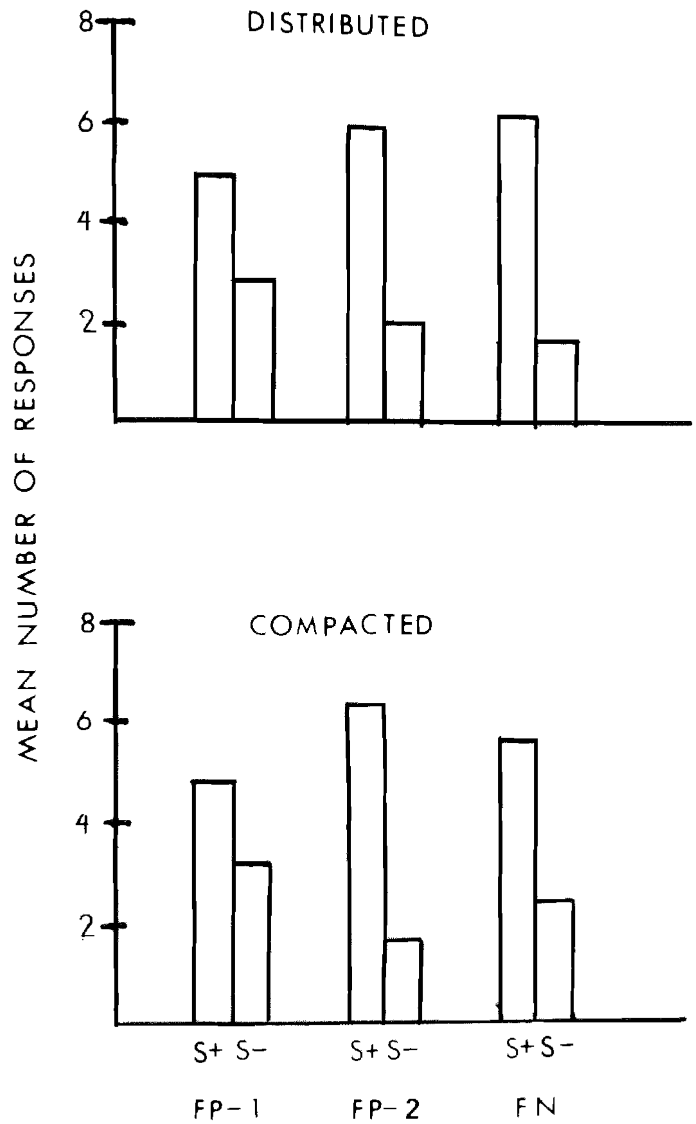

Fig. 10. Bar graph of original pretraining stimulus generalization test results.

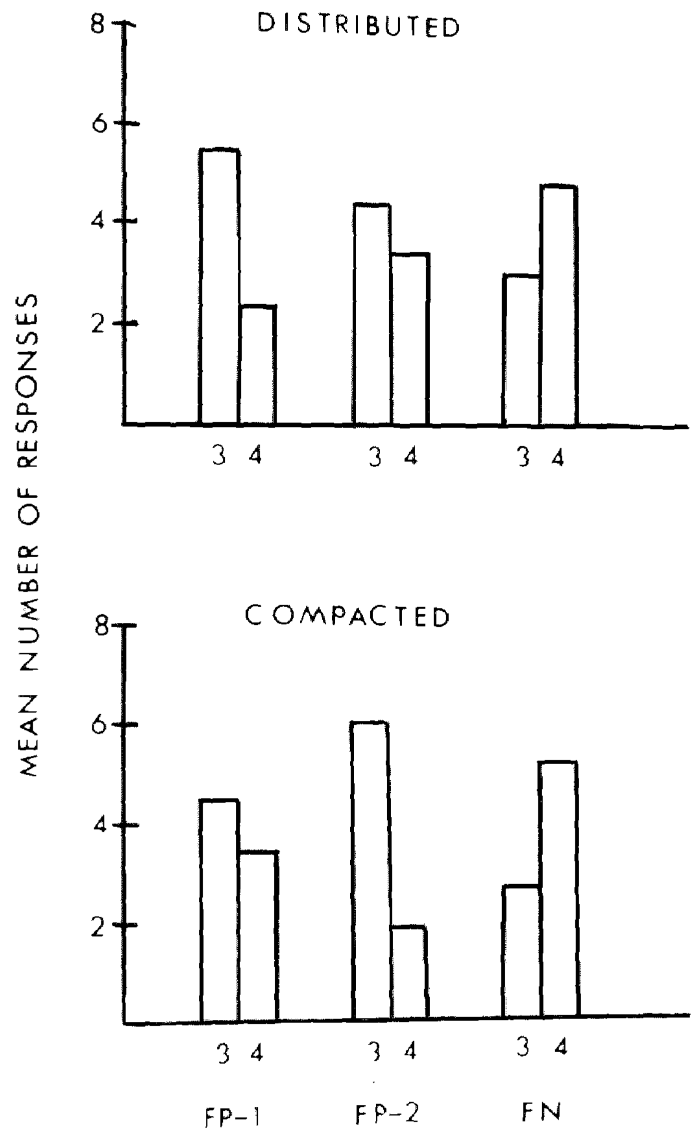

Fig. 11. Bar graph of 3 versus 4 elements generalization test results.

CHAPTER IV

DISCUSSION

Significantly more errors occurred when the elements were located in the outer corners of the display cards than when the elements were located near the centers of the display cards. In the distributed elements condition, a conditional discrimination failed to develop in either the FP-2 or FN conditions. The children failed to modify their responses to _A_ based on the presence or absence of _B_. The children did not have to make a conditional discrimination in the FP-1 condition. They needed only to respond to _B_ to learn. Most of the children who solved the discrimination responded exclusively to _B_ in the last ten trials. The pigeons in the studies of Jenkins and Sainsbury (1969, 1970) did not solve the discrimination until responses to _A_ had ceased and only _B_ responses were emitted. Both organisms used a similar response pattern in solving the discrimination.

Compacting the elements eliminated the feature-positive effect and significantly reduced FP-2 errors. The FP-2 compacted group made significantly fewer errors than the FP-2 distributed group but significantly more errors than the FP-1 and FN compacted groups. Sainsbury (1971b, 1973) reported that those few children who did learn in the FN condition appeared to use _B_ as a cue to respond to the other display. These children would appear to search out B with their finger without touching B and then make a remote response to the other display. A remote response strategy could successfully solve both the FP-2 and FN discriminations. Murphy and Miller (1959) and Jeffrey and Cohen (1964) state that a remote response strategy is difficult to learn because of the cue-response separation. They also suggest that the difficulty of learning when cue and response are separated is a function of the distance between cue and response. The findings of this study as well as those of Campione and Beaton (1972) do not support cue-response difficulty.

Sainsbury (1971b) suggests that failure to learn in the FN condition might be attributed to the difficulty in acquiring a remote response. If the acquisition of a remote response is influenced by the distance between cue and response, compacting the elements should not have changed the difficulty of the FN discrimination because the separation between cue and response remains approximately the same. Instead, compacting the elements facilitated the discrimination. In the FP-2 condition, the child needed only to locate B and then make a remote response to an A element on the same display. Compacting the display made the distance between cue and response much shorter than that in the FN compacted elements condition. However, FP-2 children made significantly more errors than the FN children in the compacted elements condition. The greater cue-response separation was in the FN case and thus it should have been the more difficult discrimination. It may have been that too few children used the remote response strategy and thus were not influenced by the independent variable of cue-response separation. Simple probability theory offers an explanation of why the FP-2 discrimination is more difficult than the FN discrimination. In the FN condition four to eight elements are positive while in the FP-2 condition only three of eight elements are positive. Since B responses are never reinforced in either condition and A responses are partially reinforced, 57% of the A responses in the FN condition are positive while only 43% of the A responses are positive in the FP-2 condition. This fact should be considered in any direct comparison of the level of difficulty of these two conditions.

The stimulus identification hypothesis is one way of integrating the feature discrimination data. The major assumption is that subjects respond either to the individual elements of the feature discrimination or to the pattern created by the arrangement of the elements. Each strategy is called a stimulus identification response (SIR). A second assumption is that the strategy remains relatively fixed through training. A theory that incorporates Gestalt principles states that compacting the elements creates patterns to which the subject can respond, and thus, the discrimination becomes a choice between two patterns created by the compacted elements. There is some evidence of pattern responding in the generalization tests.

Generalization test results show that some children appear to use an elemental SIR, some use a pattern SIR, and some use a combination of both. In the FP-1 condition (see Appendix D), S1 appears to have learned an elemental discrimination in the "REV" test and a pattern discrimination in the "NEW" and "3 v. 4" tests while S6 appears to respond on the basis of pattern exclusively. In the FP-2 condition (see Appendix E), S17 appears to have learned solely on the basis of pattern, whereas, S19 appears to have responded on the basis of both element and pattern. In the FN condition (see Appendix F), S5 appears to have responded on the basis of element, S8 on the basis of pattern, and S1 on the basis of both element and pattern. Sainsbury (1971b) found that for his children, common elements alone had exerted little control over responding and that a 3 v. 4 common elements generalization test resulted in no differential responding. The results of this more recent study indicate that the children's responses were affected not only by the common element itself, but also by the number of common elements in the 3 v. 4 generalization test.

Schuck (1960) has demonstrated that monkeys tend to limit the scope of their observations to the area of the manipulanda. Young children are also thought to limit the scope of their observations to the immediate area of their finger when emitting a finger pointing response. One possible explanation of why compacting elements facilitates FN learning is that B now remains in sight toward the end of A responses on the AB card. This may allow the inhibitory strength of B to suppress AB responses. No reduction of responding to B was observed in generalization tests for the compacted elements group in Sainsbury's 1971a study. Children in the FN compacted group in this study did respond differentially to individual A and B elements in the original pretraining elements generalization test. This would indicate that individual elements were noticed. These results should be interpreted with caution, however, because such effects could be attributed to differential pretraining.

According to the simultaneous discrimination model the organism responds to the individual elements rather than the display as a whole, leaving each element (common or distinctive) with its own reinforcement probability. The FP-2 group results indicated that the children responded to the distinctive display because the distinctive feature signaled reinforcement. Responses do not have to be made to the distinctive feature to gain stimulus control as originally hypothesized by Jenkins and Sainsbury (1969, 1970).

Feature positive children appear to learn a general response to a display with a distinctive element, since responses generalize to other displays with common and distinctive elements other than those originally used. Response to the distinctive display generalized even when responses to the distinctive feature were not reinforced as in the FP-2 condition. Children in the FN condition appear to learn on the basis of both number and type of element. In both the distributed and compacted elements groups the FN children preferred the four common element card over the three common element card in the three versus four common elements generalization test and the number of responses to the distinctive feature increased when common and distinctive elements were interchanged in the reversed elements generalization test.

The proximity of elements appears to be a critical veriable in determining feature discrimination difficulty. However, to what extent remote responding influences that difficulty is unclear. Several studies have found that learning a discrimination is inversely related to the distance between cue and response (c.f., Murphy & Miller, 1955, 1958, 1959; Jeffrey & Cohen, 1964). Sainsbury (1971b) suggests that remote responding may explain the difficulty of acquiring a feature-negative discrimination. For the feature-negative subjects in this study, a remote response is shorter in the distributed than in the compacted elements condition. These results either contradict the assumption that remote response difficulty is a function of distance or suggest that a factor other than cue-response separation determined feature-negative difficulty. Results of the feature-positive two condition support the hypothesis that learning is adversely affected by larger distances between cue and response.

Compacting the elements to the centers of the displays improved the design of Sainsbury's (1971a) experiment by providing more information about the subjects responses to the compacted displays during actual training trials by allowing responses to the individual elements to be recorded. This was not possible in Sainsbury's study. The results showed that in all the compacted elements conditions, common and distinctive elements were responded to differentially. These subjects appeared to be responding to the individual elements as well as to a pattern created by the elements.

The generalization data showed that children appeared to respond to the displays on the basis of element or pattern or both. The particular strategy used by the children appeared to be independent of the condition to which they were assigned. These results need to be interpreted with caution. The specific directions to respond to an element may have prompted an elemental response. Further investigations should examine the effects of verbal instructions on responding.

Future research should also investigate the function of the characteristics and complexity of the stimulus displays. Independent variables that can be investigated include type, number, size, redundance, similarity, saliancy, and proximity of elements as well as the function of quadrants separating the elements. Dependent variables to investigate may include latency of responding. Change-over of responding could be investigated in a free operant situation. As more dimensions are added and displays become more complex, it will be necessary to investigate how the subject responds in the presence of each of these dimensions. A developmental question then arises. That is, do subjects of different ages respond differently to the displays. Finally feature discrimination results can be compared to current theories to learning to see how they account for the results.

Reference Note

1. Bitgood, S. C., Segrave, K., & Jenkins, H. M. Effects of negative verbal feedback on the feature positive effect. Manuscript submitted for publication 1974.

References

Broadbent, D. E. Perception and communication. New York: Pergamon, 1958.

Campione, J. C., & Beaton, V. L. Transfer of training: Some boundary conditions and initial theory. Journal of Experimental Child Psychology, 1972, 13, 94-114.

D'Amato, M. R., & Schiff, D. Overlearning and brightness-discrimination reversal, Journal of Experimental Psychology, 1965, 69, 375-381.

Ehrenfreund, D. An experimental test of the continuity theory of discrimination learning with pattern vision. Journal of Comparative Physiological Psychology, 1948, 41, 408-422.

Goodwin, W. R., & Lawrence, D. H. The functional independence of two discrimination habits associated with a constant stimulus situation. Journal of Comparative and Physiological Psychology, 1955, 48, 437-443.

Hearst, E. Excitation, inhibition, and discrimination learning. In N. J. Mackintosh and W. K. Honig (Eds.), Fundamental issues in associative learning. Halifax: Dalhousie University Press, 1969, 1-41.

Hilgard, E. R., & Bower, G. H. Theories of learning. New York: Appleton-Century-Crofts, 1966.

Hill, W. F. Learning: A survey of psychological interpretations. Scranton: Chandler Publications, 1971.

Hull, C. L. Principles of Behavior. New York: Appleton-Century-Crofts, 1943.

Hull, C. L. A behavior system. New Haven, Connecticut: Yale University Press, 1952.

Jeffrey, W. E., & Cohen, L. B. Effect of spatial separation of stimulus response and reinforcement in selective learning in children. Journal of Experimental Psychology, 1964, 67, 577-580.

Jenkins, H. M. Noticing and responding in a discrimination based on a distinguishing element. Learning and motivation, 1973, 4, 115-137.

Jenkins, H. M., & Sainsbury, R. S. The development of stimulus control through differential reinforcement. In N. J. Mackintosh and W. K. Honig, (Eds.), Fundamental issues in associative learning. Halifax: Dalhousie University Press, 1969, 123-167.

Jenkins, H. M., & Sainsbury, R. S. Discrimination learning with the distinctive feature on positive or negative trials. In D. Mostofsky, (Ed.), Attention: Contemporary theory and analysis. New York: Appleton-Century-Crofts, 1970, 239-275.

Kendler, H. H., & Kendler, T. S. Vertical and horizontal processes in problem solving. Psychological Review, 1962, 69, 1-16.

Kimble, G. A. Hilgard and Marquis' conditioning and learning. New York: Appleton-Century-Crofts, 1961.

Lashley, K. S. An examination of the "contiguity theory" as applied to discriminative learning. Journal of General Psychology, 1942, 26, 241-265.

Lawrence, D. H. Acquired distinctiveness of cues: I, Transfer between discriminations on the basis of familiarity with the stimulus. Journal of Experimental Psychology, 1949, 39, 770-784.

Lawrence, D. H. The nature of a stimulus: Some relationships between learning and perception. In S. Koch (Ed.), Psychology: A Study of a Science (Volume V). New York: McGraw-Hill, 1963.

Lawrence, D. H., & DeRivera, J. Evidence for relational discrimination. Journal of Comparative Physiological Psychology, 1954, 47, 465-471.

Lovejoy, E. P. An analysis of the overlearning reversal effect. Psychological Review, 1966, 73, 87-103.

Mackintosh, N. J. The effect of attention on the slope of generalization gradients. British Journal of Psychology, 1965a, 56, 87-93.

Mackintosh, N. J. Selective attention in animal discrimination learning. Psychological Bulletin, 1965b, 64, 124-150.

McClearn, G. E., & Harlow, H. F. The effect of spatial contiguity on discrimination learning by rhesus monkeys. Journal of Comparative and Physiological Psychology. 1954, 47, 391-394.

Meyer, D. R., Polidora, V. J., & McConnell, D. G. Effects of spatial S-R contiguity and response delay upon discriminative performances by monkeys. Journal of Comparative and Physiological Psychology, 1961, 54, 175-177.

Miller, R. E., & Murphy, J. V. Influence of the spatial relationships between the cue, reward, and response in discrimination learning. Journal of Experimental Psychology, 1964, 67, 120-123.

Murphy, J. V., & Miller, R. E. The effect of spatial contiguity of cue and reward in the object-quality learning of rhesus monkeys. Journal of Comparative and Physiological Psychology, 1955, 48, 221-229.

Murphy, J. V., & Miller, R. E. The effect of the spatial relationship between the cue, reward, and response in simple discrimination learning. Journal of Experimental Psychology, 1958, 56, 26-31.

Murphy, J. V., & Miller, R. E. Spatial contiguity of cue, reward and response in discrimination learning by children. Journal of Experimental Psychology, 1959, 58, 485-489.

Norton, G. R., Muldrew, D., & Strub, H. Feature-positive effect in children. Psychonomic Science, 1971, 23, 317-318.

Pilek, V. The feature positive effect as a function of changing stimulus structure. Unpublished master's thesis. University of Calgary, 1972.

Pubols, B. H. The facilitation of visual and spatial discrimination reversal by overtraining. Journal of Comparative Physiological Psychology, 1956, 49, 243-248.

Reid, L. S. The development of noncontinuity behavior through continuity learning. Journal of Experimental Psychology, 1953, 46, 107-112.

Rescorla, R. A., & Wagner, A. R. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black and W. F. Prokasy (Eds.), Classical conditioning II: Current research and theory. New York: Appleton-Century-Crofts, 1972.

Restle, F. The selection of strategies in cue learning. Psychological Review, 1962, 63, 329-343.

Sainsbury, R. S. Discrimination learning utilizing positive or negative cues. Canadian Journal of Psychology, 1973, 27, 46-57.

Sainsbury, R. S. Effect of proximity of elements on the feature positive effect. Journal of the Experimental Analysis of Behavior, 1971a, 16, 315-326.

Sainsbury, R. S. The "feature positive effect" and simultaneous discrimination learning. Journal of Experimental Child Psychology, 1971b, 11, 347-356.

Schuck, J. R. Pattern discrimination and visual sampling by the monkey. Journal of Comparative and Physiological Psychology, 1960, 53, 251-255.

Spence, K. W. Analysis of formation of visual discrimination habits in the chimpanzee. Journal of Comparative Psychology, 1937a, 23, 77-100.

Spence, K. W. The differential response in animals to stimuli varying within a single dimension. Psychological Review, 1937b, 44, 430-444.

Spence, K. W. The nature of the response in discrimination learning. Psychological Review, 1952, 59, 89-93.

Sperling, S. E. Reversal learning and resistance to extinction. Psychological Bulletin, 1965a, 63, 281-297.

Sperling, S. E. Reversal learning and resistance to extinction: A supplementary report. Psychological Bulletin, 1965b, 64, 310-312.

Stollnitz, F., & Schrier, A. M. Discrimination learning by monkeys with spatial separation of cue and response. Journal of Comparative and Physiological Psychology, 1962, 55, 876-881.

Suchman, R. G., & Trabasso, T. Stimulus preference and cue function in young children's concept attainment. Journal of Experimental Child Psychology, 1966, 3, 188-198.

Sutherland, N. S. Stimulus analyzing mechanisms. In Proceedings of a symposium on the mechanization of thought processes (Volume 2). London: H.M.S.O., 1959.

Sutherland, N. S., & Mackintosh, N. J. Discrimination learning: Non additivity of cues. Nature, 1964, 201, 528-530.

Trabasso, T. R., & Bower, G. H. Attention in learning: Theory and research. New York: John Wiley and Sons, 1968.

Wyckoff, L. B. The role of observing responses in discrimination learning. Part I. Psychological Review, 1952, 59, 431-442.

Zeaman, D., & House, B. J. The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), Handbook of mental deficiency. New York: McGraw-Hill, 1963.

APPENDIX A

FP-1 ERRORS PER BLOCKS OF TEN TRIALS

			B1	B2	B3	B4	Total
(C1) Distributed	Circle (D1)	S1	4	5	1	3	13
		S2	2	0	0	0	2
		S3	0	0	0	0	0
		S4	1	0	0	0	1
		S5	0	0	0	0	0
			7	5	1	3	16
	Square (D2)	S6	7	1	8	1	17
		S7	0	0	0	0	0
		S8	0	0	0	0	0
		S9	0	0	0	0	0
		S10	2	0	0	0	2
			9	1	8	1	19
(C2) Compacted	Circle (D1)	S11	5	5	5	5	20
		S12	0	0	0	0	0
		S13	0	0	0	0	0
		S14	3	4	3	4	14
		S15	1	0	0	0	1
			9	9	8	9	35
	Square (D2)	S16	3	0	0	0	3
		S17	5	2	3	0	10
		S18	4	0	0	0	4
		S19	1	0	0	0	1
		S20	6	0	0	0	6
			19	2	3	0	24

APPENDIX B

FP-2 ERRORS PER BLOCKS OF TEN TRIALS

Condition	Group	Subject	B1	B2	B3	B4	Total
(C1) Distributed	Circle (D1)	S1	3	2	0	0	5
		S2	6	8	5	5	24
		S3	2	0	0	0	2
		S4	8	3	0	0	11
		S5	7	7	6	1	21
			26	20	11	6	63
	Square (D2)	S6	6	10	0	0	16
		S7	8	8	5	6	27
		S8	7	8	7	7	29
		S9	7	7	5	5	24
		S10	7	6	9	5	27
			35	39	26	23	123
(C2) Compacted	Circle (D1)	S11	8	7	7	5	27
		S12	2	0	0	0	2
		S13	7	6	5	6	24
		S14	3	0	0	0	3
		S15	5	8	0	0	13
			25	21	12	11	69
	Square (D2)	S16	3	0	0	0	3
		S17	4	5	5	0	14
		S18	3	0	0	0	3
		S19	1	0	0	0	1
		S20	6	6	0	0	12
			17	11	5	0	33

APPENDIX C

FN ERRORS PER BLOCKS OF TEN TRIALS

			B1	B2	B3	B4	Total
(C1) Distributed	Circle (D1)	S1	2	1	2	0	5
		S2	1	0	0	0	1
		S3	2	0	0	0	2
		S4	7	7	3	8	25
		S5	5	4	5	4	18
			17	12	10	12	51
	Square (D2)	S6	2	0	0	0	2
		S7	2	0	0	0	2
		S8	5	4	5	2	16
		S9	2	0	0	0	2
		S10	7	3	4	5	19
			18	7	9	7	41
(C2) Compacted	Circle (D1)	S11	4	3	2	1	10
		S12	1	0	0	0	1
		S13	0	0	0	0	0
		S14	2	0	0	0	2
		S15	0	0	0	0	0
			7	3	2	1	13
	Square (D2)	S16	2	2	3	1	8
		S17	0	0	0	0	0
		S18	1	0	0	0	1
		S19	1	0	0	0	1
		S20	1	0	0	0	1
			5	2	3	1	11

APPENDIX D

FP-1 GENERALIZATION TEST

			OTS			REV			NEW			OPT		3v.4		
		TE*	A	B	/A	A	B	/A	A	B	/A	O	S	3	4	B**
Circle (D1)	S1	0	0	8	0	1	0	7	0	8	0	8	0	8	0	8
	S2	1	5	1	2	4	2	2	4	2	2	5	3	4	4	
	S3	0	2	5	1	2	5	1	1	6	1	2	6	4	4	
	S4	2	3	4	1	5	1	2	2	3	3	5	3	5	3	
	S5	13	2	0	6	3	2	3	3	1	4	4	4	4	4	
		16	12	18	10	15	10	15	10	20	10	24	16	25	15	

(C1) Distributed

Square (D2)	S6	0	0	8	0	1	7	0	0	8	0	0	8	8	0	
	S7	0	1	6	1	1	4	3	2	4	2	2	6	5	3	
	S8	0	0	8	0	4	0	4	4	0	4	0	8	6	2	
	S9	2	1	5	2	3	1	4	2	2	4	3	5	4	4	
	S10	17	5	1	2	0	8	0	0	8	0	8	0	8	0	
		19	7	28	5	9	20	11	8	22	10	13	27	31	9	

Circle (D1)	S11	0	0	8	0	1	7	0	0	7	1	8	0	7	1	1
	S12	14	3	0	5	0	6	2	2	2	4	1	7	5	3	
	S13	1	5	3	0	6	2	0	2	5	1	8	0	0	8	
	S14	0	0	8	0	4	0	4	0	8	0	8	0	5	3	
	S15	20	1	7	0	1	4	3	2	4	2	6	2	5	3	
		35	9	26	5	12	19	9	6	26	8	31	9	22	18	

(C2) Compacted

Square (D2)	S16	1	8	0	0	1	7	0	1	7	0	8	0	7	1	
	S17	6	3	0	5	4	1	3	3	2	3	8	0	4	4	
	S18	10	3	4	1	3	3	2	4	3	1	2	6	6	2	
	S19	3	2	5	1	3	2	3	4	1	3	2	6	4	4	
	S20	4	3	1	4	4	0	4	5	0	3	3	5	3	5	
		24	19	10	11	15	13	12	17	13	10	23	17	24	16	

*TE Total Errors for S
**B Blank sector of Three common element card

APPENDIX E

FP-2 GENERALIZATION TEST

			OTS			REV			NEW			OPT		3v.4	
		TE	A	B	/A	A	B	/A	A	B	/A	O	S	3	4
	S1	24	1	1	6	2	2	4	3	3	2	2	6	3	5
Circle	S2	2	2	6	0	0	8	0	0	8	0	0	8	7	1
(D1)	S3	11	7	0	1	0	8	0	0	8	0	0	8	7	1
	S4	21	7	0	1	0	8	0	2	0	6	0	8	5	3
	S5	5	6	0	2	4	0	4	4	0	4	0	8	1	7
		63	23	7	10	6	26	8	9	19	12	2	38	23	17
(C1) Distributed															
	S6	27	0	4	4	2	3	3	4	2	2	5	3	3	5
	S7	29	4	0	4	2	1	5	4	0	4	5	3	3	5
Square	S8	24	4	2	2	2	3	3	3	1	4	3	5	3	5
(D2)	S9	27	3	0	5	2	6	0	2	3	3	8	0	5	3
	S10	16	2	6	0	0	8	0	0	8	0	0	8	8	0
		123	13	12	15	8	21	11	13	14	13	21	19	22	18
	S11	2	1	7	1	0	8	0	0	8	0	3	5	7	1
Circle	S12	24	5	1	2	3	3	2	2	0	6	0	8	7	1
(D1)	S13	3	8	0	0	0	8	0	0	7	1	0	8	8	0
	S14	13	6	1	1	2	5	1	4	3	1	0	8	7	1
	S15	27	4	1	3	3	2	3	3	1	4	5	3	2	6
		69	24	10	6	8	26	6	9	19	12	8	32	31	9
(C2) Compacted															
	S16	14	2	6	0	2	1	5	2	3	3	0	8	4	4
	S17	3	8	0	0	8	0	0	8	0	0	8	0	8	0
Square	S18	1	7	0	1	0	8	0	0	8	0	8	0	5	3
(D2)	S19	12	8	0	0	0	8	0	0	8	0	8	0	8	0
	S20	3	6	0	2	0	8	0	0	8	0	8	0	5	3
		33	31	6	3	10	25	5	10	27	3	32	8	30	10

APPENDIX F

FN GENERALIZATION TEST

		TE	OTS A	OTS B	OTS /A	REV A	REV B	REV /A	NEW A	NEW B	NEW /A	OPT O	OPT S	3v.4 3	3v.4 4
Circle (D1)	S1	2	0	0	8	0	8	0	1	0	7	0	8	0	8
	S2	25	5	2	1	3	2	3	2	2	4	3	5	2	6
	S3	18	4	1	3	2	4	2	4	3	1	7	1	7	1
	S4	1	2	0	6	1	6	1	0	7	1	1	7	4	4
	S5	5	2	0	6	1	7	0	4	0	4	0	8	4	4
		51	13	3	24	7	27	6	11	12	17	11	29	17	23
(C1) Distributed															
Square (D2)	S6	2	2	0	6	1	7	0	1	7	0	8	0	4	4
	S7	16	2	3	3	0	2	6	1	4	3	4	4	6	2
	S8	2	0	0	8	0	0	8	0	0	8	8	0	1	7
	S9	19	3	0	5	4	4	0	4	2	2	6	2	3	5
	S10	2	0	0	8	0	0	8	0	2	6	8	0	0	8
		41	7	3	30	5	13	22	6	15	19	34	6	14	26
Circle (D1)	S11	0	3	1	4	3	1	4	2	1	5	3	5	4	4
	S12	2	0	0	8	0	1	7	0	0	8	7	1	2	6
	S13	1	5	1	2	0	6	2	0	0	8	0	8	2	6
	S14	0	0	0	8	0	8	0	0	5	3	0	8	0	8
	S15	10	4	1	3	2	2	4	2	0	6	4	4	4	4
		13	12	3	25	5	18	17	4	6	30	14	26	12	28
(C2) Compacted															
Square (D2)	S16	0	0	0	8	1	2	5	1	0	7	7	1	0	8
	S17	1	2	1	5	2	1	5	4	0	4	5	3	2	6
	S18	1	1	0	7	0	3	5	0	0	8	8	0	1	7
	S19	1	3	0	5	0	8	0	0	8	0	8	0	6	2
	S20	8	1	1	6	4	0	4	1	3	4	2	6	6	2
		11	7	2	31	7	14	19	6	11	23	30	10	15	25

APPENDIX G

B/AB RESPONSES FOR TRIALS 1-10

			FP-1	FP-2	FN
(C1) Distributed	Circle (D1)	S1	2/6	1/8	0/2
		S2	8/8	0/4	0/1
		S3	10/10	1/9	0/2
		S4	7/9	1/3	3/7
		S5	10/10	2/5	0/5
	Square (D2)	S6	0/3	1/5	1/2
		S7	10/10	2/4	0/2
		S8	10/10	2/5	2/5
		S9	10/10	4/7	0/2
		S10	8/8	3/6	2/7
(C2) Compacted	Circle (D1)	S11	4/5	3/5	3/4
		S12	10/10	1/9	0/1
		S13	10/10	1/4	0/0
		S14	1/7	1/7	1/2
		S15	3/9	0/5	0/0
	Square (D2)	S16	6/7	0/7	0/2
		S17	5/5	0/6	0/0
		S18	6/6	0/8	0/1
		S19	0/9	0/9	1/1
		S20	0/4	2/6	0/1

APPENDIX H

RATIO OF B-RESPONSES TO TOTAL TRAINING TRIALS

			FP-1	FP-2	FN
(C1) Distributed	Circle (D1)	S1	9/40	1/22	0/39
		S2	11/13	0/40	0/12
		S3	10/10	1/12	0/14
		S4	8/12	1/25	5/40
		S5	10/10	2/40	0/40
	Square (D2)	S6	0/40	2/30	1/13
		S7	10/10	5/40	2/18
		S8	10/10	3/40	9/40
		S9	10/10	7/40	2/12
		S10	11/13	4/40	7/40
(C2) Compacted	Circle (D1)	S11	19/40	9/40	5/40
		S12	10/10	1/13	0/12
		S13	10/10	3/12	0/10
		S14	2/40	1/15	1/14
		S15	4/11	0/30	0/10
	Square (D2)	S16	16/20	0/15	1/40
		S17	6/39	2/37	0/10
		S18	10/14	0/13	1/11
		S19	1/11	0/11	1/17
		S20	0/19	3/26	1/12